JANUARY 01 1976

Formalism for the rotation matrix of rotations about an arbitrary axis

Alan Palazzolo

American Journal of Physics 44, 63–67 (1976)
https://doi.org/10.1119/1.10140

Formalism for the rotation matrix of rotations about an arbitrary axis

Alan Palazzolo
Department of Physics and Astronomy, The University of Toledo, Toledo, Ohio 43606
(Received 23 April 1975; revised 18 June 1975)

In this paper a formalism is developed that enables one to write down by inspection the matrix which represents the transformation between two coordinate systems, one of which is rotated by an angle θ about an arbitrary axis. Only the direction cosines of the rotation axis and θ need be known. It is pointed out that this transformation is not discussed in the usual treatments of classical mechanics.

INTRODUCTION

Traditional methods of representing coordinate transformations are somewhat limited in their range of applications. In standard treatment of rotations,[1,2] the student learns that he can multiply rotation matrices to obtain a single matrix, which represents a transformation that is equivalent to a series of rotations. Unfortunately, the student's feeling of competence in rotational transformations can be turned to one of inadequacy on encountering certain examples of rotations about a noncoordinate axis. He is then challenged with the formidable task of calculating the angles of rotations about coordinate axes, equivalent to the desired rotation. An example that illustrates this difficulty is found in Marion's[1] text. His problem (problem 1-1) asks the student to rotate a rectangular coordinate system by 120° about an axis making equal angles with each of the coordinate axes (the crystallographic 1,1,1 axis). As has been pointed out by Kittel,[3] this problem is easily solved by noticing the symmetry of the system: for a 120° rotation, $\hat{e}_x$ becomes $\hat{e}_z'$, $\hat{e}_y$ becomes $\hat{e}_x'$, and $\hat{e}_z$ becomes $\hat{e}_y'$, where the primes denote the rotated coordinate frame. What would have happened, though, if the rotation had been 119°?

On reexamining the usual formalism, it begins to emerge that the source of the problem is that all rotation matrices presented to the student are for rotations about one of the coordinate axes and this problem simply cannot be done with the use of such rotations. It should be remarked at this point that an apparently reasonable approach is first to use two successive rotations to align the $\hat{e}_z$ axis with the 1,1,1 direction and then to rotate by 120° (or any other angle in the general case). Clearly, this method fails, since after these rotations the $\hat{e}_z'$ is along the old 1,1,1 direction rather than the $\hat{e}_x$ axis, as we know from symmetry it must be.

The formalism for rotations about the coordinate axes is seen to be essentially equivalent to the Euler angles, a series of rotations about the coordinate axes of the initial frame as well as intermediate frames. In order to solve the problem of rotations about a noncoordinate axis, another formalism is required. Such a formalism is developed in this paper, and it will be seen that it is possible to write down a matrix representing a rotation through any angle about an arbitrary axis passing through the origin. In the case that the rotation axis coincides with a coordinate axis, the more usual rotation matrices are of course reproduced.

Formally, if a prime denotes the rotated system, the transformation between systems is

$$x_i' = \sum_j \lambda_{ij} x_j ,$$

where

$$\lambda_{ij} \equiv \cos(x_i', x_j),$$

with the notation (x_i', x_j) meaning the angle between unit vectors in the x_i' direction and in the x_j direction. The derivation of a formalism for the direction cosines λ_{ij}, which follows, is founded on geometric arguments. The formalism is shown to be consistent with the orthogonality condition

$$\sum_j \lambda_{ij} \lambda_{ij} = \delta_{ii} .$$

EXPLANATION OF FIG. 1

In Fig. 1, plane N is defined to be perpendicular to the rotation axis, **R**, which is line $O\bar{O}$. Plane N is at unit distance from the origin. The orientation angles, α, β, and γ describe the position of **R** relative to the original frame, and are measured from the positive $\hat{X}_1$, $\hat{X}_2$, and $\hat{X}_3$ axes, respectively. The quantities a, b, and c represent the distances along the $\hat{X}_1$, $\hat{X}_2$, and $\hat{X}_3$ axes to the intersection of plane N with the original frame. The new frame is represented by axes, $\hat{X}_1'$, $\hat{X}_2'$, and $\hat{X}_3'$, similarly, the quantities a', b', and c' are the distances along these axes to plane N. The angle θ is the rotation angle about R. The rotation axis passes through the origin and $\bar{O}$, which is the center of rotation in plane N.

DEFINING LENGTHS IN FIG. 1

If $\hat{e}_r$ is the unit vector along the rotation axis R,

$$\hat{e}_r = \cos\alpha\,\hat{e}_1 + \cos\beta\,\hat{e}_2 + \cos\gamma\,\hat{e}_3 .$$

Define (see Fig. 1) the vectors

$$\mathbf{V}_{AB} = a\hat{e}_1 - b\hat{e}_2 ,$$

$$\mathbf{V}_{CA} = -a\hat{e}_1 + c\hat{e}_3 ,$$

$$\mathbf{V}_{BC} = b\hat{e}_2 - c\hat{e}_3 .$$

Since N is perpendicular to R,

$$\mathbf{V}_{AB} \cdot \hat{e}_r = \mathbf{V}_{CA} \cdot \hat{e}_r = \mathbf{V}_{BC} \cdot \hat{e}_r = 0.$$

Since $\bar{O}$ is in plane N,

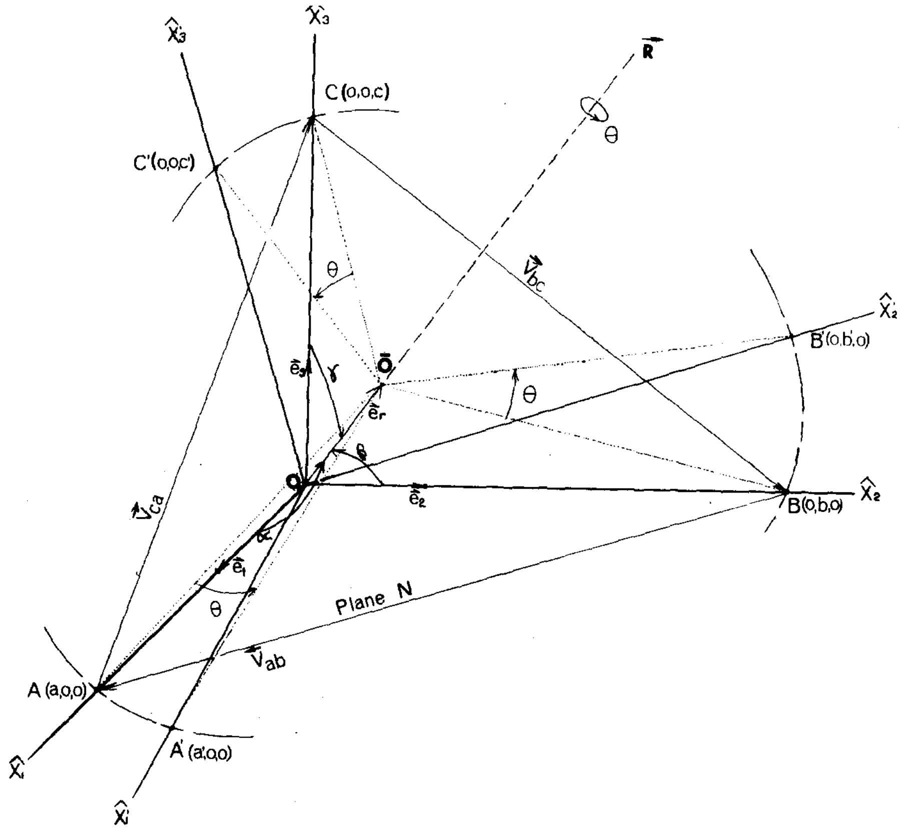

Fig. 1. Illustration of the basic parameters involved in the derivation. Line $O\bar{O}$ is the axis of rotation, **R**, and plane N is orthogonal to **R**.

$$(\mathbf{OA} - \mathbf{O\bar{O}}) \cdot (\mathbf{O\bar{O}}) = 0,$$

where $\mathbf{O\bar{O}} = \hat{e}_r$. From these conditions, three equations can be written and solved for a, b, and c. They are

$$(a\hat{e}_1 - b\hat{e}_2) \cdot (\cos\alpha\hat{e}_1 + \cos\beta\hat{e}_2 + \cos\gamma\hat{e}_3) = 0, \quad (1)$$

$$(-a\hat{e}_1 + c\hat{e}_3) \cdot (\cos\alpha\hat{e}_1 + \cos\beta\hat{e}_2 + \cos\gamma\hat{e}_3) = 0, \quad (2)$$

$$[(a - \cos\alpha)\hat{e}_1 - \cos\beta\hat{e}_2 - \cos\gamma\hat{e}_3] \cdot (\cos\alpha\hat{e}_1 + \cos\beta\hat{e}_2 + \cos\gamma\hat{e}_3) = 0, \quad (3)$$

or

$$a\cos\alpha - b\cos\beta = 0,$$
$$-a\cos\alpha + c\cos\gamma = 0,$$
$$a\cos\alpha = \cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1,$$

from which

$$a = \sec\alpha, \quad (4)$$
$$b = \sec\beta, \quad (5)$$
$$c = \sec\gamma. \quad (6)$$

Note that

$$a = 0a = 0a',$$
$$b = 0b = 0b',$$
$$c = 0c = 0c'.$$

Note that segment lengths are preserved since points A, B, and C stay in plane N during rotation about R (see Fig. 2).

From Fig. 2 and Eq. (4), note that

$$\bar{0}a = (1 + \sec^2\alpha - 2\sec\alpha\cos\alpha)^{1/2}$$
$$= \tan\alpha$$
$$= \bar{0}a'. \quad (7)$$

Likewise,

$$\bar{0}b = \tan\beta = \bar{0}b' \quad (8)$$

and

$$\bar{0}c = \tan\gamma = \bar{0}c'. \quad (9)$$

These segments remain in plane N during rotation about R, so their lengths are preserved also.

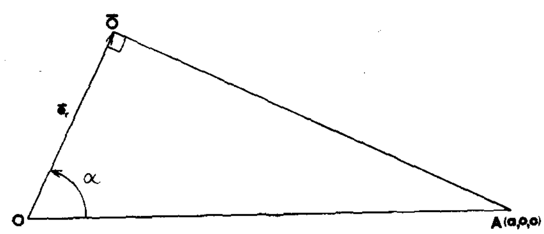

Fig. 2. The unit vector $\hat{e}_r$ is normal to plane N, which contains line segment $\bar{O}A$.

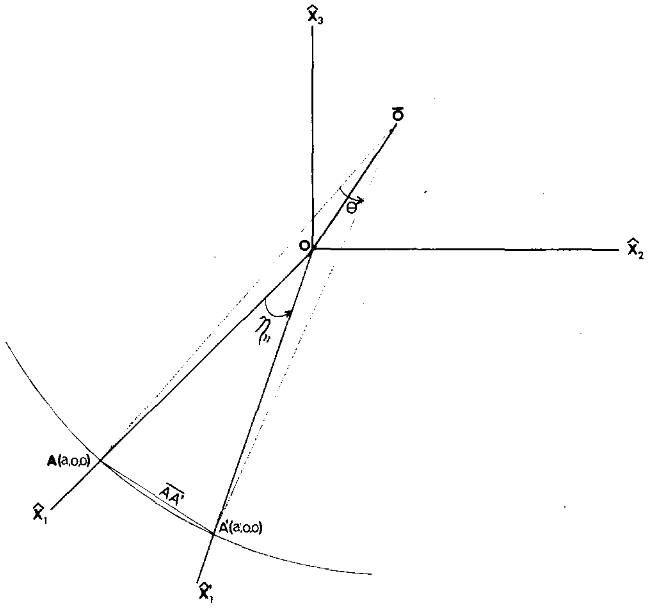

Fig. 3. In deriving the diagonal elements, triangles $\bar{O}AA'$ and OAA' are of key importance. Note that angle $\eta_{11} = (\hat{X}_1', \hat{X}_1) \neq \theta$.

FORMALISM FOR THE DIAGONAL ELEMENTS OF THE TRANSFORMATION MATRIX

From Fig. 3, consider triangle $\bar{O}AA'$. By the law of cosines,

$$AA'^2 = 2(\bar{0}a)^2 - 2(\bar{0}a)^2 \cos\theta,$$

since $\bar{O}a = \bar{O}a'$. From Eq. (7),

$$AA'^2 = 2\tan^2\alpha - 2\tan^2\alpha \cos\theta.$$

Next, consider triangle OAA', for which (again from the law of cosines)

$$\cos\eta_{11} = \frac{2(0a)^2 - (AA')^2}{2(0a)^2}$$

$$= \frac{2\sec^2\alpha - 2\tan^2\alpha(1-\cos\theta)}{2\sec^2\alpha}$$

$$= 1 - \sin^2\alpha + \sin^2\alpha\cos\theta$$

$$= \cos^2\alpha + \sin^2\alpha\cos\theta.$$

By definition, $\cos\eta_{11} = \lambda_{11}$.

Constructing similar diagrams between the $\hat{X}_2$ and $\hat{X}_2'$ axes and between the $\hat{X}_3$ and $\hat{X}_3'$ axes, λ_{22} and λ_{33} are found to be

$$\lambda_{22} = \cos^2\beta + \sin^2\beta\cos\theta$$

and

$$\lambda_{33} = \cos^2\gamma + \sin^2\gamma\cos\theta.$$

CALCULATION OF OFF-DIAGONAL ELEMENTS

From Fig. 4, consider triangle OAB. Using Eqs. (4) and (5) yields

$$\overline{AB} = (OA^2 + OB^2)^{1/2} = (\sec^2\alpha + \sec^2\beta)^{1/2}.$$

Considering triangle $\bar{O}AB$ and using the law of cosines, note that

$$\cos\zeta_{12} = \frac{\bar{0}a^2 + \bar{0}b^2 - \overline{AB}^2}{2\bar{0}a\bar{0}b}$$

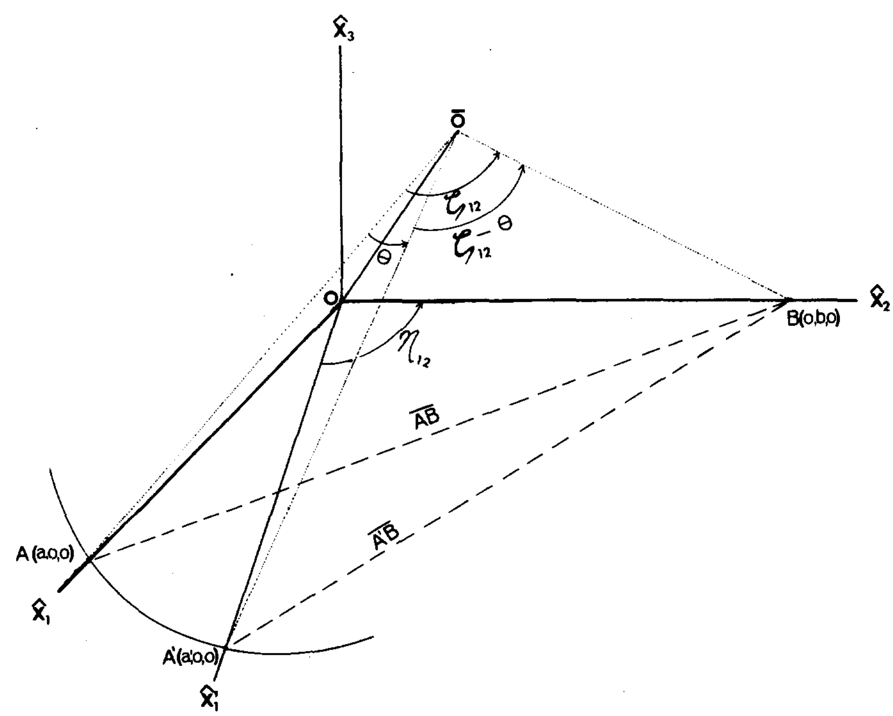

Fig. 4. After rotation through angle θ about $O\bar{O}$, the cosine of the angle between $\hat{X}_1'$ and $\hat{X}_2$ is equal to λ_{12}. Note that lengths from origin to intersection of coordinate axes with plane N are unaltered.

$$= \frac{\tan^2\alpha + \tan^2\beta - \sec^2\alpha - \sec^2\beta}{2\tan\alpha\tan\beta}$$

Using the trigonometric identity $\tan^2\psi - \sec^2\psi = -1$ yields

$$\cos\zeta_{12} = -\cot\alpha\cot\beta$$

and

$$\sin\zeta_{12} = (1 - \cot^2\alpha\cot^2\beta)^{1/2}.$$

From triangle $\bar{O}A'B$, we see

$$\begin{aligned}\overline{A'B}^2 &= [\bar{0}a'^2 + \bar{0}b^2 - 2\bar{0}a'\bar{0}b\cos(\zeta_{12} - \theta)] \\ &= \bar{0}a'^2 + \bar{0}b^2 - 2\bar{0}a'\bar{0}b(\cos\zeta_{12}\cos\theta + \sin\zeta_{12}\sin\theta) \\ &= \tan^2\alpha + \tan^2\beta \\ &\quad - 2\tan\alpha\tan\beta[-\cot\alpha\cot\beta\cos\theta \\ &\quad + (1 - \cot^2\alpha\cot^2\beta)^{1/2}\sin\theta] \\ &= \tan^2\alpha + \tan^2\beta + 2\cos\theta \\ &\quad - 2\sin\theta(\tan^2\alpha\tan^2\beta - 1)^{1/2}.\end{aligned}$$

Finally, consider triangle $OA'B$.

$$\begin{aligned}\cos\eta_{12} &= \cos(\hat{X}_1', \hat{X}_2) = \lambda_{12} \\ &= \frac{OA'^2 + OB^2 - \overline{A'B}^2}{2OA'OB} \\ &= [\sec^2\alpha + \sec^2\beta - \tan^2\alpha - \tan^2\beta \\ &\quad - 2\cos\theta + 2\sin\theta(\tan^2\alpha\tan^2\beta - 1)^{1/2}] \\ &\quad \times (2\sec\alpha\sec\beta)^{-1}.\end{aligned}$$

Using the identity $\sec^2\psi - \tan^2\psi = 1$ yields

$$\begin{aligned}\lambda_{12} &= \frac{1 - \cos\theta + \sin\theta(\tan^2\alpha\tan^2\beta - 1)^{1/2}}{\sec\alpha\sec\beta} \\ &= \cos\alpha\cos\beta - \cos\alpha\cos\beta\cos\theta \\ &\quad + \sin\theta(\sin^2\alpha\sin^2\beta - \cos^2\alpha\cos^2\beta)^{1/2}.\end{aligned}$$

Using

$$\begin{aligned}\sin^2\alpha\sin^2\beta &= (1 - \cos^2\alpha)(1 - \cos^2\beta) \\ &= 1 - \cos^2\beta - \cos^2\alpha + \cos^2\alpha\cos^2\beta\end{aligned}$$

yields

$$\begin{aligned}\lambda_{12} &= \cos\alpha\cos\beta - \cos\alpha\cos\beta\cos\theta \\ &\quad + \sin\theta(1 - \cos^2\beta - \cos^2\alpha)^{1/2}.\end{aligned}$$

Since $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$,

$$\lambda_{12} = \cos\alpha\cos\beta - \cos\alpha\cos\beta\cos\theta + \sin\theta\cos\gamma.$$

Performing an identical geometric analysis to find the remaining direction cosines, that is, $\cos\eta_{ij} = \cos(\hat{X}_i', \hat{X}_j)$, the following results are obtained:

$$\lambda_{11} = \cos^2\alpha + \sin^2\alpha\cos\theta, \qquad (10)$$

$$\lambda_{12} = \cos\alpha\cos\beta - \cos\alpha\cos\beta\cos\theta + \sin\theta\cos\gamma, \qquad (11)$$

$$\lambda_{13} = \cos\alpha\cos\gamma - \cos\alpha\cos\gamma\cos\theta - \sin\theta\cos\beta, \qquad (12)$$

$$\lambda_{21} = \cos\beta\cos\alpha - \cos\beta\cos\alpha\cos\theta - \sin\theta\cos\gamma, \qquad (13)$$

$$\lambda_{22} = \cos^2\beta + \sin^2\beta\cos\theta, \qquad (14)$$

$$\lambda_{23} = \cos\beta\cos\gamma - \cos\beta\cos\gamma\cos\theta + \sin\theta\cos\alpha, \qquad (15)$$

$$\lambda_{31} = \cos\gamma\cos\alpha - \cos\gamma\cos\alpha\cos\theta + \sin\theta\cos\beta, \qquad (16)$$

$$\lambda_{32} = \cos\gamma\cos\beta - \cos\gamma\cos\beta\cos\theta - \sin\theta\cos\alpha, \qquad (17)$$

$$\lambda_{33} = \cos^2\gamma + \sin^2\gamma\cos\theta. \qquad (18)$$

These equations can be expressed more concisely as

$$\begin{aligned}\lambda_{ij} &= \cos\sigma_i\cos\sigma_j \\ &\quad + [\epsilon_{ijk}\sin\theta\cos\sigma_k - \cos\sigma_i\cos\sigma_j\cos\theta](1 - \delta_{ij}) \\ &\quad + \delta_{ij}(\sin^2\sigma_i\cos\theta), \qquad (19)\end{aligned}$$

where k is a dummy index such that $i = k = j$, and $\alpha = \sigma_1$, $\beta = \sigma_2$, and $\gamma = \sigma_3$. Also here δ_{ij} = the Kronecker delta symbol, and ϵ_{ijk} = the Levi–Cevita density. Here we take θ as positive when measured counterclockwise.

AN EXAMPLE: SOLUTION TO MARION'S PROBLEM 1-1

For the problem suggested earlier, namely, rotation by 120° about an axis making equal angles with the original coordinate axes, $\sigma_1 = \sigma_2 = \sigma_3 = \cos^{-1}(3^{-1/2})$ and $\theta = 120°$. From Eq. (19) we find $\lambda_{11} = \lambda_{13} = \lambda_{21} = \lambda_{22} = \lambda_{32} = \lambda_{33} = 0$ and $\lambda_{12} = \lambda_{23} = \lambda_{31} = 1$ or

$$\begin{bmatrix} X_1' \\ X_2' \\ X_3' \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix},$$

which is the obvious solution.

It is also quite simple to show that the trace of the transformation matrix represented by Eq. (19) is $1 + 2\cos\theta$, as Goldstein[4] shows that it must be.

ACKNOWLEDGMENTS

The author wishes to thank Dr. William F. Rush for his assistance in the development of this paper, and Suzanne Strall for her patience in typing the manuscript.

APPENDIX: PROOF OF THE ORTHOGONALITY CONDITION

To prove that Eq. (19) represents an orthogonal transformation, substitute the equation into the relation at the end of the Introduction. The orthogonality condition now states

$$\sum_j \lambda_{ij}\lambda_{lj} = \sum_j [\cos\sigma_i \cos\sigma_j + (\epsilon_{ijk}\sin\theta\cos\sigma_k - \cos\sigma_i\cos\sigma_j\cos\theta) + \delta_{ij}(\cos\theta\sin^2\sigma_i - \epsilon_{ijk}\sin\theta\cos\sigma_k + \cos\sigma_i\cos\sigma_j\cos\theta)]$$
$$\times [\cos\sigma_l\cos\sigma_j + (\epsilon_{ljm}\sin\theta\cos\sigma_m - \cos\sigma_l\cos\sigma_j\cos\theta) + \delta_{lj}(\cos\theta\sin^2\sigma_l - \epsilon_{ljm}\sin\theta\cos\sigma_m + \cos\sigma_l\cos\sigma_j\cos\theta)].$$

This expression can be expanded into 36 terms, of which many combine to equal zero, or are exactly equal to zero. The following types of terms arise in the expansion and are evaluated in the following manner.

(i) A recurring term in the expansion is of the form

$$\sum_j \cos\sigma_j \cos\sigma_k \epsilon_{ijk}.$$

This sum consists of two equal and opposite nonzero terms, due to the reversal in signs of the Levi–Cevita symbol. In the summation over j, j and k alternate between the two remaining non-i values, according to the condition of Eq. (19) that $i \neq k \neq j$. This term is always equal to zero. The same results hold if i is replaced by l and k by m.

(ii) Another term in the expansion is of the form

$$\sum_j \delta_{ij}\delta_{lj}\epsilon_{ljm} F(\sigma_i, \sigma_j, \sigma_k, \ldots).$$

The Kronecker delta symbol, δ_{lj}, is nonzero if j is equal to l, but in this case the Levi–Cevita symbol equals zero. This term then remains zero in all cases.

(iii) The expression

$$\epsilon_{ilk}\cos\sigma_k + \epsilon_{lim}\cos\sigma_m$$

occurs only once in the expansion. The first term is generated under the condition that $j \equiv l$ (because it is part of a term multiplied by δ_{lj} and summed over j). Then by the condition on Eq. (19) k cannot equal i or l. The second term is generated under the condition that $j \equiv i$ (for the same reason as the first term, with δ_{lj} replaced by δ_{ij}). Similarly, m cannot be equal to i or l. Since all indices are restricted to 1, 2, or 3, the only case that can occur is $k \equiv m$. In this case the expression is equal to zero because $\epsilon_{ilk} = -\epsilon_{lik}$.

By utilizing these results, the orthogonality condition becomes

$$\sum_j \lambda_{ij}\lambda_{lj} = \cos\sigma_i\cos\sigma_l + \cos^2\theta\Big(-\cos\sigma_i\cos\sigma_l + \sum_j \delta_{ij}\delta_{lj}(\sin^2\sigma_i\sin^2\sigma_l + \cos\sigma_l\sin^2\sigma_i\cos\sigma_j + \cos\sigma_i\sin^2\sigma_l\cos\sigma_j + \cos\sigma_i\cos\sigma_l\cos^2\sigma_j)\Big) + \sin^2\theta\sum_j \cos\sigma_k\cos\sigma_m\epsilon_{ijk}\epsilon_{ljm}. \qquad \text{(A.1)}$$

Consider $i = l$ for one case of the orthogonality condition. Then

$$\sum_j \lambda_{ij}\lambda_{ij} = \cos^2\sigma_i + \cos^2\theta(1 - \cos^2\sigma_i) + \sin^2\theta\sum_j \cos^2\sigma_k\epsilon_{ijk}^2.$$

The coefficient of $\sin^2\theta$ results because $k \equiv m$ for the sum to be nonzero. The sum can be carried out over j and produces the results

$$\sum_j \cos^2\sigma_k\epsilon_{ijk}^2 = \sum_j \cos^2\sigma_j - \cos^2\sigma_i = 1 - \cos^2\sigma_i.$$

This follows because k alternates between both remaining indices (not including i) during the summation. Then, in the case $i = l$, the sum $\Sigma_j\lambda_{ij}^2 = 1$.

Next consider Eq. (A.1) for the case where $i \neq l$. The term $\Sigma_j\delta_{ij}\delta_{lj}(\cdots)$ is zero by definition of the Kronecker deltas and the condition $i \neq l$. The orthogonality condition then reduces to

$$\sum_j \lambda_{ij}\lambda_{lj} = \cos\sigma_i\cos\sigma_l + \cos^2\theta[-\cos\sigma_i\cos\sigma_l] + \sin^2\theta[\sum_j \cos\sigma_k\cos\sigma_m\epsilon_{ijk}\epsilon_{ljm}].$$

Consider the coefficient of $\sin^2\theta$. Since $i \neq l$, j can only assume one value to have a nonzero term in the summation. Also, to keep all terms of the Levi–Cevita symbols distinct, $k \equiv l$ and $m \equiv i$. The coefficient then becomes

$$\sum_j \cos\sigma_l\cos\sigma_i\epsilon_{ijl}\epsilon_{lji} = -\cos\sigma_i\cos\sigma_l\sum_j \epsilon_{ijl}^2 = -\cos\sigma_l\cos\sigma_i.$$

By substituting this result into the orthogonality condition, the desired results are obtained; that is, $\Sigma_j\lambda_{ij}\lambda_{lj} = 0$, $i \neq l$.

[1]J. B. Marion, *Classical Dynamics of Particles and Systems* (Academic, New York, 1970), 2nd ed.

[2]K. R. Symon, *Mechanics* (Addison-Wesley, Reading, MA, 1960), 2nd ed.

[3]C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1971), 4th ed., p. 141.

[4]H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, MA, 1950), 1st ed., p. 123.